DOCUMENT RESUME

ED 173 350 TM 007 608

AUTHOR Welch, Wayne W.
TITLE Goal-Free Formative Evaluation--An Example.
PUB DATE Mar 78
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978); Best copy available

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Check Lists; *Evaluation Criteria; *Evaluation Methods; Evaluators; *Formative Evaluation; Negative Attitudes; Program Evaluation; *Publications; Summative Evaluation; *Textbook Evaluation
IDENTIFIERS *Goal Free Evaluation

ABSTRACT

A formative goal-free evaluation was conducted to assess products (documents) developed after the first year of a federally funded project. Fifteen packets of sixty-six documents were rated by an external panel of five judges, using a checklist developed for the evaluation. An intermediary between the panel and the goal-free evaluator interviewed key project staff to clarify questions about the documents and to obtain perceptions of the goal-free process. The bulk of the written evaluation report was a summary of judges' comments and ratings for each packet. Several problems surfaced: project staff seemed anxious and mistrusting about the goal-free method, and there was a major discrepancy between proposed project goals and the external evaluator's perception of the project. In summary, goal-free evaluation seems most appropriate when product development is a major task. Personal observation during evaluation seriously jeopardizes the disclosure of goals. The approach is best used as an adjunct to other evaluation methods, although it needs refining. (The checklist is appended). (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED173350

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

GOAL-FREE FORMATIVE EVALUATION--AN EXAMPLE

Wayne W. Welch

University of Minnesota

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Wayne W. Welch

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

March, 1978

Paper presented at the Annual Meeting of the American Educational Research Association.

March 27-31, 1978. Toronto, Canada.

TM007 608

BACKGROUND

In 1975, the internal evaluator of a federally funded project asked me to serve as an external evaluation consultant. His perception of this role was not well-formed at the time and he asked me to consider how such an evaluation consultant might be helpful to his project.

The situation was this. The project was nearing the end of the first year of an anticipated five-year funding pattern. The proposal had promised an internal evaluation which apparently was in operation. But the internal evaluator was concerned about several things; the possible biasing effects of his role as a half-time project administrator; the credibility of the local evaluation effort, particularly as it affected future funding probabilities; and a desire to learn more about various evaluation alternatives.

The problem had several ingredients which made it eligible for an experimental goal-free evaluation (Scriven, 1972). First, and foremost, our initial one-hour discussion had not addressed the objectives of the project except to mention the college had received a five-year grant. In fact, I was unaware of the source or amount of funding. Second, there was considerable interest on the part of the staff to support a formative evaluation of the first year's activities as a supplement to their own design. Finally, a number of products had been developed and the staff was interested in external assessment of these materials.

The author wishes to thank Ms. Sharon Studer who helped in the design of this study and served as the intermediary.

The situation seemed to meet the major criteria for qualifying as a limited formative goal-free evaluation: i.e., there was very little pre-study rhetoric on goals, the staff was interested in possible side effects of their materials, and trial versions of several products were available for review.

A subsequent meeting was held with the project director to explain the goal-free philosophy and her approval was obtained. The letter of agreement describes the goal-free approach as we envisioned it at that time.

> My evaluation responsibility would be to carry out a goal-free evaluation of products and activities completed to date. The goal-free evaluation would be designed to assess the products from the standpoint of a student and/or parent considering the program. It would include materials review, observation of planned activities, and possibly interviews of key administrators. A checklist would be used to rate the products and activities.
>
> The goal-free approach is an experimental venture which capitalizes upon objectivity, reduces personal bias, and increases potential for detecting side effects. It also can be converted to goal-based evaluation should that become desirable.

The estimated time for the study was eight days, with about $400 being allocated for attendant expenses: travel, typing, assistant's salary, and the like. However, there were more documents than I had anticipated, and the total time was 11 days for the evaluator and about 120 hours of graduate assistant help. The total cost for the evaluation was about $2000.

PROCEDURES

Goal-free methodology was not well established in 1975, but some techniques needed to be selected that would not violate the intent of the goal-free approach. There was also a severe constraint on the resources available for the study. We eventually decided to gather information in two

ways: an external review panel would rate the products, and an intermediary would interview key project personnel. She, in turn, would be interrogated by the goal-free evaluator. However, the intermediary would carefully censure the evaluator's questions so as not to divulge the project goals and only would answer questions related to current project activities and concerns. The key caution lights were questions about goals and questions that required answers in the past or future tense.

All documents developed by project funding were submitted to an external review panel who were knowledgeable in evaluation but unaware of the project or its objectives. The documents were evaluated using a checklist developed for this project since none were found in the literature. The checklist addressed those components generally considered important for effective products (see Alkin and Fink, 1974) and provided opportunities for judges to consider potential side effects. An overall quality rating was also requested of the judges. (The checklist is found in Appendix 1.)

Each packet of documents was reviewed by five judges. The judges included the author of this report, a graduate student majoring in evaluation (who also was the intermediary), and students drawn from an introductory evaluation course. The packets of documents (15 packets which contained 66 documents) were distributed across the students so that each document was reviewed by three student judges. (It is of interest to note that the average rating time per packet was 62 minutes, somewhat longer than expected.) The fourth judge completed checklists for each packet while the fifth judge synthesized the ratings into Evaluation Summaries for each packet.

The intermediary interviewed six key members of the project to clarify questions related to the products being judged, to obtain perceptions of the goal-free process, and most important, to make sure the evaluation was in the right ball park. (There was an ever-present feeling on my part that I may have been missing some critical aspect of the study. It is somewhat like setting up a blind procedure to evaluate a liquid you think is wine but want to be assured occasionally it is not, in fact, floor wax. The intermediary and the internal evaluator helped to make sure we were not setting up a procedure to taste floor wax.)

Two important components in the panel rating process created some problems and should be mentioned here. First, the total number of documents to be rated was greater than we realized, and it required considerable effort on the part of the staff to gather these together in a useful way. Eventually (after a month's delay) the 66 documents were categorized into 15 packets by the internal evaluator. (I learned later these match roughly some administrative structure within the project.) This cooperation by the internal staff is a key element in the goal-free methodology.

Second, the task of synthesizing and reporting the four ratings per packet was difficult. Yet, a summary of reactions seemed more useful to the project than forwarding all of the judges' checklists. Eventually, it was decided to prepare a written report which contained an Evaluation Summary for each packet. This report was forwarded to the project administration for their perusal. Then, a meeting was held with the internal evaluator to answer any

questions he might have on the report. However, we avoided crossing the goal-free threshold until he met with the project administration to determine their wishes.

A meeting was then held with the six project coordinators, the internal evaluator, and the goal-free team. The first hour was spent interviewing the team for further information, and then a dialogue was held where the project objectives were revealed and we looked for discrepancies between the intended objectives and the goal-free perceptions.

RESULTS

The bulk of the written report was a summary of the judges' comments and ratings for each packet. First, the documents in each packet were identified. This was followed by a guess at the intended purpose of the product. The judges' comments were divided into two groups: perceived strengths and perceived limitations. Finally, a general quality grade was given to the documents which was the mean of the four ratings. A letter grade was given by the judges, and the quality grade is the mean of those grades. The ratings permitted a quality comparison among the 15 packets.

A sample Evaluation Summary follows.

I. Product D.1, D.2, D.3, D.4

Four manuals for occupational therapy assistants; labeled (1) Basic Client Support Skills, (2) Physical Dysfunction, (3) Developmental Disabilities/Mental Retardation, and (4) Organization and Administration of Activity Departments in Long-Term Care Facilities.

II. Apparent Purpose

The purposes of the manuals are stated in the introduction; to provide students with the skills to work in long-term care centers. The target students are those enrolling in an occupational therapy assistant program. The manuals are designed to help students achieve specific behavioral objectives.

III. Comments

A. Strengths

1. Material is clearly presented and easy to understand.
2. The introduction in three of the four manuals helps facilitate understanding of these documents.
3. I would guess that these materials would work, although there is no evidence of this.
4. Low cost reproduction.

B. Limitations

1. Evaluation criteria for achieving objectives should be discussed with students. How will objectives be measured?
2. Repetitive format may cause boredom among students.
3. The unit is very much oriented toward vocabulary--that is, learning the medical terms--e.g., thermoplast for silly putty, edema for swelling. Certainly, the vocabulary must create enormous problems for students. This should be considered in the revisions.
4. Seems to ignore the problem of the students' attitudes toward patients with long-term afflictions--fears, anxieties, etc. Focuses instead on the cognitive aspects of treatment.
5. Much knowledge will be developed, but seems short on application. Will students be able to use their knowledge in actual settings?
6. Major points in journal articles should be highlighted.
7. Some use of unidentified acronyms--e.g., COTA, OTR.
8. Document D.4 less successful than the first three. Use of behavioral objectives here seems less than successful.
9. Often assumes patient is "he" and therapist is "she".
10. Need to know how these fit into total program.

IV. Quality

Documents	Ratings	Sum	Mean	Grade
D.1	A-, A, B+, B-	37	9.25	B+
D.2	A-, B, B, B-	33	8.25	B
D.3	A-, B+, B+, B-	35	8.75	B+
D.4	C, B+, B+, B-	30	7.50	B-/B
		105	8.75	B+

The report also included a set of general impressions by the goal-free evaluator and a summary of the site interviews conducted by the intermediary.

Some examples of impressions were such things as:

> The quantity of work produced in the short time the project has been in existence (I believe 1-1/2 years) is impressive. I expected perhaps 10-20 products; the actual total was 66.
>
> The use of behavioral objectives has much merit in organizing and planning instruction; however, it is not without its limitations. (For example, judges mentioned boredom, concentration of trivial content, and stifling of creativity.) The approach should be used with caution. They must be treated as a servant, not your master.

The report was intended as a starting point for discussions between the staff developers and the evaluation team, both internal and external. As a formative study, our main intent was to highlight those features of the products that "caught our eye" as outside evaluators. We hoped these observations would be considered carefully in product revisions.

REFLECTIONS

One purpose of this study was to learn more about the goal-free approach. How did it work in practice? In what ways did it differ from other evaluations I've done? What are some of its strengths and limitations? Several things emerged in this study that might be useful to those considering the approach.

The anxiety level of the staff seemed very high during this study. The staff was concerned about a formless evaluator judging their work without even talking to them. Perhaps more importantly, they felt the products alone did not give a full picture of the project. And although the internal evaluator explained the goal-free evaluation was only one aspect of the total evaluation program, they generally were uncomfortable with methodology.

Perhaps this feeling can be illustrated by noting the concern over my initial choice of an intermediary. His wife, a nurse, had taught at the institution

for one year and had decided teaching was not her desire. She resigned, quite amicably, as far as I know. However, the staff was worried that this association might influence his judgment and questioned my choice on several occasions. Finally, in frustration and anger, my intermediary quit and I had to select another.

I think the major problem (or perhaps strength) here is the lack of opportunity for a feeling of mutual trust to be established between the staff and the external evaluator. Without interacting personally with me to better know my motives, style, interests, or qualifications, the threat of the evaluation seemed greater.

Several aspects of the study concerned me as the evaluator. The scope of the study was too small, even with free graduate student help, to give the approach an adequate test. The checklist, though judged adequate by the panel, needs revision and refinement. I was constantly worried that I was missing large parts of the project and felt quite helpless to do anything about it. The intermediary interviews were designed to help check on major omissions, but my interrogation of her suggested important process components were being ignored (for example, watching the students in action). Unfortunately, the limited resources did not permit us to pursue this very far.

Several staff argue we should compare their accomplishments with what they started with, not with some idealized set of standards. This is especially a problem when new activities were being established, e.g., alumni newsletter. Selecting a fair standard is a difficult task for the goal-free evaluator, especially when he is so ignorant of project antecedents. The issue needs further exploration, I believe.

A major discrepancy between the project rhetoric and my perception of the project was the enormous attention given to behavioral objectives. The complete proposal was written in bureaucratic style with major goals, sub-project goals, and specific program objectives. That was a surprise revelation to me from examining products alone. A few documents contained a set of objectives, but the theme could not be detected by examining such things as faculty development brochures or alumni publications.

It seems likely the rhetoric of proposal writing does not always carry through to project implementation. But in this case, I think a more general phenomenon was occurring. The few personal contacts I had with staff members suggested a style of operation that seemed contrary to the structure implied in the behavioral approach. The staff seemed more like humanists than engineers. Perhaps this influenced my perceptions. The effects of the few personal interactions may have taken on great importance. Although reduction of staff influence is a key purpose of goal-free evaluation, the personal factor seemed heightened in the goal-free mode. This was especially true for the intermediary. (Of course we must be cautious that what really was happening was an increased awareness of the powerful influence of the personal contacts in a regular evaluation.)

Several aspects of the evaluation seemed to work well. My estimates of the resource allocation to the several sub-projects based on the quantity and quality of the sub-project products was quite accurate. That is, it was possible to compare the amount spent on faculty development with the amount allocated to student recruitment once the various categories were identified for me. This gives some validity to the study. Also, several unintended side effects were identified to the project staff: for example, vocabulary problems

in student materials, occasional sexist words or pictures, and a patronizing attitude toward faculty development. However, for the most part, the panel's ratings were high and the strengths outnumbered the limitations.

This last point prompts me to point out the need for caution in judging products in this way. The evaluator must be careful not to become too infatuated by the occasional unanticipated or negative outcome. These findings, though the grist of the anecdotal mill, must properly be balanced against the many positive aspects of the products. Seeking positive side effects is equally important in the goal-free approach.

In summary, I would offer the following recommendations for those considering a goal-free evaluation. It seems most appropriate when the development of products is clearly the major task of the program. To use goal-free evaluation when a great deal of personal observation of the process is necessary seriously jeopardizes the disclosure of goals. When this occurs, it's no longer a goal-free evaluation.

Goal-free evaluation is probably best used as an adjunct to other kinds of evaluations. It highlights the need for objectivity, even if it doesn't guarantee it, and it stimulates both staff and evaluator to work hard at their task.

The approach is in need of improved methodology. The present study was quite modest and only was able to try a few techniques. My general impression is that the value slightly outweighed the limitations, particularly as a stimulus for heightened interest in evaluation. I believe additional applications of goal-free evaluation are needed to refine our methods and to give the goal-free approach a fair and adequate test.

References

Scriven, Michael. "Prose and Cons about Goal-Free Evaluation." Evaluation Comment, Vol. 3, No. 4 (December, 1972) pp. 1-4.

Alkin, Marvin C. and Fink, Arlene. "Evaluation Within the Context of Product Development: A User Orientation." In Evaluating Educational Programs and Products (Gary D. Borich, ed.) Educational Technology Publications, Englewood Cliffs, N.J. (1974).

Appendix 1

GOAL-FREE CHECKLIST

1. Brief description of the document/activity. (Include name, developer, distribution means, and physical attributes.)

2. Content. (Summarize in one paragraph.)

3. Apparent intended purpose/objectives. (What can be learned from the materials?)

4. Apparent audience/required entry behavior. (What background is required for user?)

5. Underlying assumptions. (List and rate explicitness.)

6. Evidence or judgment of how developed and tested. (What was done? What evidence of success?)

7. Evidence or judgment of product effectiveness. (Consider alternative ways to achieve the same objectives.)

8. Evidence or judgment of product efficiency. (Estimate benefit for cost or time required.)

9. Potential side effects.

10. Overall quality rating. (Consider clarity, integrity, conciseness. Rate on an A - F scale as you might grade a class project.)

Product Code: Reviewer's Name: